

MINERvA Conditions Data Delivery

Igor Mandrichenko

MINERvA Infrastructure meeting

11/4/2010

Requirements

- We still do not know:
 - Number of jobs per hour
 - Average
 - Peak
 - Security requirements
 - Expected request correlation characteristics
 - How many jobs will request same data at the same time
 - Expectation:
 - High for data production
 - Lower for analysis

Requirements

- What we do know
 - Amount of data “chunk” for pedestals
 - 32,000 channels
 - ~10 floating point numbers per channel (conservatively 8 bytes per number)
 - ~2.5 MB total
 - All channels are recorded for every measurement
 - Conditions data is represented as a function
time -> 32K*10 quantities
 - Normally, the measurement validity interval is open-ended
 - From time of measurement on to infinity
 - Currently DB has ~1 year worth of data, 500 – 1000 IOVs per channel

Postgres and simplified schema

```
create table calibration_iovs
(
    iov_id bigserial primary key,
    begin_time timestamp
);

create index iov_begin_time_inx on iovs(begin_time);

create table calibration
(
    iov_id bigint references calibration_iovs(iov_id) on delete cascade,
    channel int default 0,
    a_min float,
    a_max float,
    a_avg float,
    b_min float,
    b_max float,
    b_avg float,
    c_min float,
    c_max float,
    c_avg float,
    primary key (iov_id, channel)
);
```

Postgres and simplified schema

- Stored 604 snapshots, 32,000 channels each
 - 604 rows in calibration_iops table
 - 19,392,000 rows in calibration table
- Insertion time: between 12 and 19 seconds per snapshot
 - Insert new row into calibration_iops table
 - Insert 32,000 rows into calibration table
- Retrieval time (give me 32,000 9-tuples for given point of time): 1 second

```
rexdb01.fnal.gov> python fetch.py "2010-01-03 05:06:07"  
32000 values for interval 2010-01-03 05:00:00 - 2010-01-03 05:10:00, fetch time=0.987528 sec
```

Conditions Data Delivery

MINERvA conditions data delivery

